

A Study on Converting bibliographic data of public libraries expressed in KORMARC into BIBFRAME

Joo-Yong Kim*, Pan-Seop Shin*

*Student, Dept. of Computer Engineering, Daejin University, Pocheon, Korea

*Professor, Dept. of Computer Engineering, Daejin University, Pocheon, Korea

[Abstract]

BIBFRAME, which is attracting attention as an alternative to the machine-readable catalog format (MARC) in the library world, presents a new bibliographic data model in the open web environment while maintaining compatibility with existing data. To convert KORMARC(Korean data model of MARC) records into BIBFRAME, we extract 25 key fields by analyzing the latest 5,000 bibliographic data from Nowon-gu Library in Seoul. The extracted core fields are classified into three types according to the compatibility of MARC 21, and define conversion rules for each type. In addition, implement an open source-based converter to perform KORMARC to BIBFRAME conversion. As a basic study on KORMARC to BIBFRAME conversion, this study is meaningful in that it analyzes the latest KORMARC information actually used, defines conversion rules, and attempts BIBFRAME conversion.

▶ **Key words:** BIBFRAME, Machine-readable cataloging, KORMARC, Conversion, Linked Open Data

[요 약]

도서관 계에서 기계 가독 목록 형식(MARC)에 대한 대안으로 주목받고 있는 BIBFRAME은 기존 데이터와의 호환성을 유지하면서 오픈 웹 환경에서 새로운 서지기술 데이터모델을 제시한다. MARC의 한국형 데이터 모델인 KORMARC 레코드의 BIBFRAME 변환을 위해, 서울시 노원구립 도서관의 최신 서지 데이터 5,000개를 분석하여 25개의 핵심 필드를 추출한다. 핵심 필드들을 MARC 21의 호환성 여부에 따라 세 가지 유형으로 분류하고, 각 유형별 변환 기법을 정의한다. 또한 오픈소스 기반의 변환기를 구현하여 KORMARC to BIBFRAME 변환 작업을 수행한다. 본 연구는 KORMARC to BIBFRAME 변환에 대한 기초연구로써, 실제 사용되는 최신 KORMARC 정보를 분석하여 변환 규칙을 정의하고, BIBFRAME 변환을 시도했다는 점에 의의가 있다.

▶ **주제어:** 비브프레임, 기계가독목록형식, 한국문헌자동화목록형식, 변환, 링크드 오픈 데이터

-
- First Author: Joo-Yong Kim, Corresponding Author: Pan-Seop Shin
 - *Joo-Yong Kim (jdragon8608@gmail.com), Dept. of Computer Engineering, Daejin University
 - *Pan-Seop Shin (psshin@daejin.ac.kr), Dept. of Computer Engineering, Daejin University
 - Received: 2021. 08. 11, Revised: 2021. 11. 22, Accepted: 2021. 11. 23.
 - This paper is an extension of the paper "Analysis of bibliographic data fields by frequency of use in public libraries" presented at the 64th Summer Conference 2021 of the Korean Society of Computer and Information.

I. Introduction

도서관에서 사용하는 서지 데이터는 종이카드에 수기로 작성하여 목록함에 배열 후 사용하는 형태에서부터 시작되었다. 이러한 형태는 신간 정보를 추가하고 폐기 도서 정보를 삭제하는 방식에 있어서 책자 형태보다 가제성이 좋아 두루 사용되었다.

이후 웹 생태계의 발전으로 정보의 대량화와 신속성이 요구되면서 도서관 서지 데이터에 대한 새로운 처리방식이 필요하게 되었다. 이러한 요구에 부응하기 위해 컴퓨터의 신속하고 정확한 데이터 처리 능력을 이용하여 전통적인 편목 업무를 보다 효율적으로 처리하고자 미국 의회 도서관에서 MARC(Machine Readable Cataloging)를 1960년대에 개발했으며, 이를 자동화 목록법 또는 기계 가독 목록법이라 한다. 오늘날 국내외 도서관마다 수십만, 수백만 건의 목록 레코드가 MARC 형식으로 저장되어 데이터베이스의 역할을 하고 있으며, 도서관 간 정보 공유의 측면에서도 MARC 형식은 여전히 중요한 수단이 되고 있다. 국내 대부분의 도서관에서 기본적으로 목록 기술 규칙을 비롯하여 이를 구조화하기 위한 MARC 형식을 사용하고 있다[1][2].

MARC의 등장으로 도서관의 서지 데이터 활용에 많은 영향을 주었지만, 오늘날의 정보환경에서는 공유와 연계를 중심으로 많은 변화가 일어나고 있어 새로운 구조의 서지 정보 기술방식이 요구되고 있다. 이에 도서관들은 MARC 형식의 서지 데이터나 전거데이터를 링크드 데이터로 변환하여 서비스하고 있다[2]. 예컨대 LC(Library of Congress, 미국 국회도서관)는 LDS(Linked Data Service)를 통해 LC가 작성한 전거데이터를 제공한다. 또한 영국 국립도서관은 'Linked Open BNB'에서 영국의 전국 서지 약 280만건을 제공하고, 프랑스 국립도서관은 전용 Open Data 포털 사이트 'data.bnf.fr'을 통해 서지 데이터와 전거데이터, 전자도서관 Gallica의 메타 데이터를 제공한다. 이외에도 독일과 스페인 국립 도서관도 기존 목록 데이터를 LOD로 변환하여 제공한다.

도서관의 또 다른 움직임은 목록 업무에 사용되는 데이터 모델로 RDF(Resource Description Framework) 기반의 Linked Data 를 사용하는 것이다. LC가 2012년 발표한 BIBFRAME (Bibliographic Framework)은 MARC 21을 대체하는 데이터 포맷의 일종이지만, 그 데이터 모델의 기반은 Linked Data를 채용하고 있다. BIBFRAME은 기존의 MARC 레코드와의 호환성을 유지하면서 MARC 를 대체할 수 있는 새로운 서지기술을 위한 데이터모델을 제시하며, 서지 정보의 기술 표준을 링크드 데이터 모델로

확장함으로써, 도서관 내부뿐만 아니라 오픈 웹 환경에서 서지 정보를 다방면으로 활용하기 위한 것이다[5].

국내에서는 BIBFRAME 적용을 위한 다양한 선행 연구들이 진행되었다. 해외의 BIBFRAME 적용 및 변환사례연구와 MARC21 과의 매핑 관련 기초연구 등이 있으며 BIBFRAME 적용을 위한 다양한 제안을 하였다. 미국에서는 미국 의회 도서관(LC)에서 제공하는 'MARC 21 to BIBFRAME conversion Tools'를 활용하여 실제 MARC 21 서지 데이터를 변환한 사례가 있다[10][13]. 아직 국내에서는 실제 변환 시도가 없어 국내 연구자들도 KORMARC 서지 데이터에 대한 BIBFRAME 매핑 및 변환 시도가 필요하다는 후속 연구의 필요성을 강조하였다 [8][9]. 이에 본 연구에서는 실제 사용 중인 KORMARC 서지정보를 분석하여 주요 필드 분석을 진행하고 추출된 필드를 기준으로 MARC 21 및 BIBFRAME 2.0 변환 명세를 참고하여 KORMARC 서지 데이터의 변환 기법을 정의한다. 또한 미국 의회도서관에서 제공하는 오픈소스를 활용하여 KORMARC 서지 데이터에 대한 BIBFRAME 2.0 변환을 수행하고자 한다. 본 연구는 BIBFRAME 간의 매핑 가능성을 도출하고 공공도서관의 서지정보를 BIBFRAME에 적용시킬 수 있는 기초 연구 자료로 활용될 수 있다.

II. Preliminaries

1. The evolution of bibliographic data

오늘날 모든 도서관에서 MARC를 이용한 자료처리가 일반화되어 있지만 미국 의회 도서관에서 처음 도서관의 전통적인 검색 수단인 목록을 컴퓨터로 처리하려는 시도를 할 때만 해도 크게 주목받지 못했다. 그렇지만 미국 의회도서관에서 1950년대 초부터 자신의 도서관 업무를 자동화하기 위한 위원회를 구성하고 업무개발을 위한 타당성을 검토하기 시작하였으며, 길버트 킹(G.M. King)을 중심으로 구성된 조사반의 권고에 따라 1960년 초부터 가능성 검토와 함께 개발 계획을 수립한 후, 1965년 MARC 1을 완성하게 됨으로써 전 세계의 도서관에 혁명적인 변화를 이끌기 시작했다. 수정 보완한 MARC 2가 1968년 제정되었는데, 이는 미국의 국가적인 표준인 USMARC Format for Bibliographic Data뿐만 아니라 세계 각국의 MARC 개발의 시초가 되었다. 캐나다의 CANMARC, 호주의 Australian MARC, 영국의 UK MARC, 독일의 MAB, 중국의 Chinese MARC, 일본의 JapanMARC, 우리나라의 KORMARC 등이 개발되었고, MARC 레코드의 국제적인


```

from xml.dom import minidom
import openpyxl

excel_file = openpyxl.Workbook()
excel_sheet = excel_file.active
excel_sheet.append(['tag', '', 'count'])

xmldoc = minidom.parse('kormarc.xml')

itemlist = xmldoc.getElementsByTagName('datafield')

wordCount = {}

for s in itemlist:
    #print('s.attributes["tag"].value.split(): ' + s.attributes["tag"].value)
    wordList = (s.attributes["tag"].value)
    wordCount[wordList] = wordCount.get(wordList, 0) + 1

keys = sorted(wordCount.keys())

for word in keys:
    print(word + ':' + str(wordCount[word]))
    excel_sheet.append([word, '', str(wordCount[word])])

excel_file.save('kormarc_tag.xlsx')
excel_file.close()
    
```

Fig. 2. Data Field Emergence Frequency Analysis Code

1.2 Analyzing frequency of use of fields

노원구립도서관 KORMARC 서지 데이터 분석 결과, 총 230개의 KORMARC Data Field 필드 중, 단행본 서지 데이터에서 사용되는 주요 필드는 25개로 나타났다. 추출된 25개의 필드는 Table. 1과 같다.

Table 1. Extracted meaningful KORMARC fields

Field	Definition	Frequency
020	국제표준도서번호	6,067 (121.34%)
049	소장사항	5,000 (100%)
056	한국십진분류기호	5,000 (100%)
090	자관 청구기호	5,000 (100%)
245	표제와 책임표시사항	5,000 (100%)
260	발행, 배포, 간사사항	5,000 (100%)
300	형태사항	5,000 (100%)
950	로컬정보-가격	5,000 (100%)
653	비통제 색인어	4,999 (99.98%)
100	기본표목-개인명	4,831 (96.62%)
507	그래픽자료의 축척 주기	4,385 (87.7%)
700	부출표목-개인명	3,287 (65.74%)
500	일반 주기	3,192 (63.84%)
740	부출표목-비통제 관련/부출 표제	2,923 (58.46%)
041	언어부호	2,328 (46.56%)
440	총서사항/부출표목-표제	1,350 (27%)
504	서지 등 주기	1,029 (20.58%)
023	출판예정도서목록제어번호	661 (12.22%)
250	판사항	575 (11.5%)
900	로컬표목-개인명	486 (9.72%)
525	부록 주기	420 (8.4%)
940	로컬표목-표제	392 (7.84%)
949	로컬표목-총서표제	194 (3.88%)
110	기본표목-단체명	169 (3.38%)
710	부출표목-단체명	79 (1.58%)

그 중, 020 필드는 중복사용이 가능하므로 출현 빈도가 121.34% 였으며, 100%인 필드는 049, 056, 090, 245, 260, 300, 950 총 7개이다. 100% 미만 ~ 90% 이상 범위의 출현 빈도로 나타난 필드는 653, 100 총 2개이며, 90%

미만 ~ 50% 이상 범위의 출현 빈도로 나타난 필드는 507, 700, 500, 740 총 4개이다. 50% 미만 ~ 10% 이상 범위의 출현 빈도로 나타난 필드는 041, 440, 504, 023, 250 총 5개이며, 10% 미만 ~ 1% 이상 범위의 출현 빈도로 나타난 필드는 900, 525, 940, 949, 110, 710 총 6개이다.

출현 빈도가 90% 이상인 필드는 도서 검색 및 자료 이용 시 필수적인 데이터라 할 수 있으며, 국내 공공 도서관의 서지 데이터를 BIBFRAME으로 전환 시 변환 기술 방법을 우선적으로 고려하여 적용시킬 필요가 있다.

2. Conversion to BIBFRAME

2.1 Classification of Fields for Mapping

BIBFRAME 2.0은 공식적으로 MARC 21에 대한 BIBFRAME 변환 명세만 존재한다. 따라서 KORMARC에 대한 BIBFRAME 변환 절차를 진행하기 위해 본 연구에서는 KORMARC의 특성을 이용하여 단계적 Mapping 절차를 진행한다.

MARC 21을 기반으로 개발된 KORMARC는 로컬데이터를 기술하는 필드 이외에 다른 필드들은 MARC 21과 상당히 유사하다. 따라서 KORMARC 사용 빈도 분석에서 추출된 필드들을 우선적으로 MARC 21의 필드들과 호환성 여부를 검토하며 이를 세 가지로 분류한다. 첫째, 완벽하게 호환이 되는 경우 'Conversion'으로 구분하며 이는 KORMARC의 서지 데이터를 BIBFRAME으로 즉시 변환할 수 있는 수준이다. 둘째, KORMARC에서는 사용하지만 MARC 21에서 존재하지 않는 필드를 구분한다. 필드 데이터의 변환이 가능하도록 국가서지 LOD를 활용하여 재정의 하며, 'Substitution'으로 구분한다.

Table 2. KORMARC Field Classification

Classification	KORMARC Field
conversion	020 / 041 / 100 / 110 / 245 / 250 / 260 / 300 / 500 / 504 / 525 / 653 / 700 / 710 / 740 /
substitution	023 / 056 / 440 / 900 / 940 / 949 / 950
no attempt to convert	049 / 090 / 507

마지막으로 KORMARC에서만 사용하거나 MARC 21에서도 사용하지만, 변환 명세가 존재하지 않는 필드는 'No attempt to convert'로 구분하고 마찬가지로 국가서지 LOD를 활용하여 변환 가능 하도록 재 정의한다. 주요 사용 필드의 분류 결과는 Table. 2 와 같다.

2.2 Defining Mapping Rules

KORMARC에서 Control Fields를 제외한 Data Fields 186개 중, 공공도서관 단행본에 주로 사용되는 필드는 25개였으며, 앞서 3분류로 정의한 필드의 매핑 분석 결과는 다음과 같다.

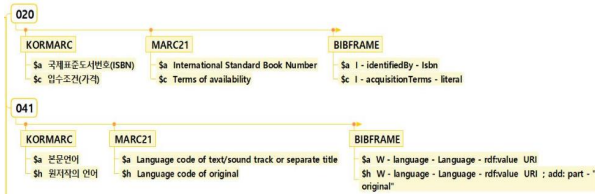


Fig. 3. 020/041 Mapping

Conversion 분류 Fig. 3의 020 필드는 I-identifiedBy-Isbn의 형태로 기술할 수 있다. 041 필드의 본문 언어(\$a)는 W-language-Language-rdf:value URI 형태로, 원저작의 언어(\$h)는 W-language-Language-rdf:URI "original" 형태로 기술할 수 있다.

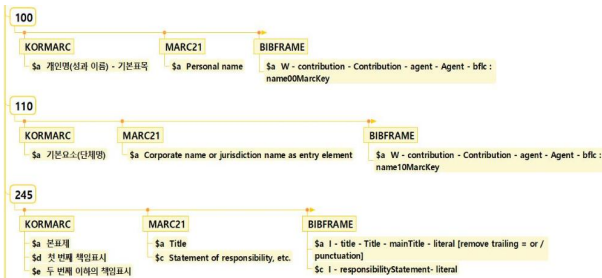


Fig. 4. 100/110/245 Mapping

Fig. 4의 100 필드는 W-contribution-Contribution-agent-Agent-bflc:name00MarcKey, 110 필드는 W-contribution-Contribution-agent-Agent-bflc:name10MarcKey, 245 필드는 I-title-Title-mainTitle-literal 형태로, 책임표시사항은 I-responsibilityStatement-literal 형태로 기술 가능하다.

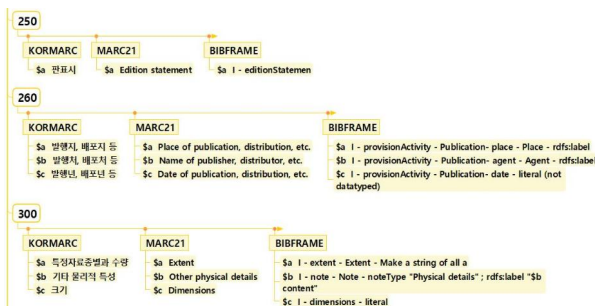


Fig. 5. 250/260/300 Mapping

Fig. 5의 250 필드는 I-editionStatement, 260 필드는 I-provisionActivity-Publication-Place-Place-rdfs:label, 발행처는 I-provisionActivity-Publication-agent-Agent-rdfs:label, 발행년은 I-provisionActivity-Publication-date-literal 형태로 변환이 가능하다. 300 필드는 I-extent-Extent-rdfs:label 형태로, 기타 물리적 특성은 I-note-Note-rdfs:label 형태로 기술한다. 크기는 I-dimensions-literal 형태로 기술한다.

Fig. 6의 500 필드는 I-note-Note-rdfs:label 형태로, 504 필드의 서지 주기 사항은 BIBFRAME 변환명세에 따라 W-supplementaryContent-SupplementaryContent-rdfs:label 형태로 기술한다. 525 필드는 부록 관련 주기 필드이며, W-supplementaryContent-SupplementaryContent 형태로 기술하며, 653 필드는 W-bf:subject-rdfs:label 형태로 기술한다.

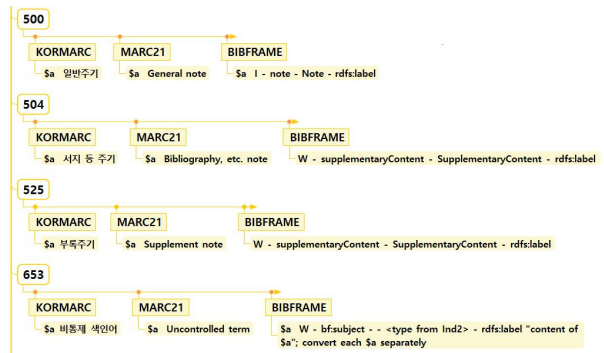


Fig. 6. 500/504/653 Mapping

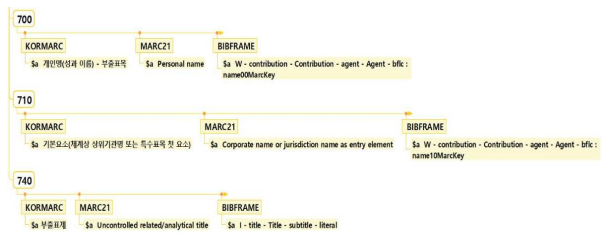


Fig. 7. 700/710/740 Mapping

700 필드는 W-contribution-Contribution-agent-Agent-bflc:name00MarcKey 형식으로 변환이 가능하다. 710 필드는 W-contribution-Contribution-agent-Agent-bflc:name10MarcKey, 740 필드는 I-title-Title-subtitle-literal 형식으로 변환한다.

Substitution 분류 Fig. 8의 023 필드는 KORMARC에서만 사용하는 필드로써 국립중앙도서관에서 부여하는 출판 예정 도서목록(CIP)에 대한 정보를 기술하는 필드이다. BIBFRAME 변환 명세가 존재하지 않아 023 필드를 변환

하기 위해서는 기술할 수 있는 LOD 정보가 필요하다. 따라서 본 연구에서는 국립중앙도서관에서 제공하는 국가서지 LOD[15][16][17] 데이터 셋을 이용하여 BIBFRAME으로 변환을 시도한다. 국가서지 LOD의 Class & Property URI 체계는 'http://lod.nl.go.kr/ontology/' 이며, KORMARC의 각 필드에 따라 사용하는 어휘를 제공하고 있다[18]. 따라서 KORMARC에서만 사용 중인 CIP 번호를 국가서지 LOD의 어휘로 대체하여 변환하면 W-Document-cip-rdf:value 형태로 기술한다.

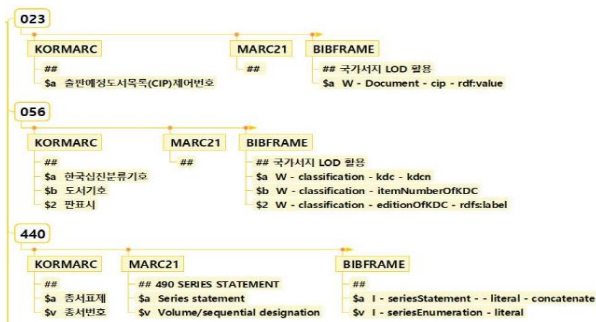


Fig. 8. 023/056/440 Mapping

056 필드는 한국십진분류기호(KDC)를 기술하는 필드로, MARC 21에서는 사용하지 않는다. 본 연구에서는 마찬가지로 국가서지 LOD를 활용하여 056 필드의 KDC 정보도 기술할 수 있도록 코드를 수정하고 BIBFRAME으로 변환을 시도한다. 056 필드의 한국십진분류기호(KDC)는 BIBFRAME에서 W-classification-kdc-kdcn 형식으로 기술 가능하며, 도서 기호는 -itemNumberOfKDC, 판표시는 -editionOfKDC-rdfs:label 형식으로 기술한다.

440 필드는 총서 사항과 총서 표제의 부출 표목이 같은 경우에 총서 사항을 기술한다. MARC 21에서는 2008년 이후 사용하지 않고, 490 필드나 830 필드에 기술하고 있다. BIBFRAME 변환 명세에 따라 총서 표제는 W-title-Title-bflc:title40MarcKey 형태로 기술하며, 총서 번호는 serisesStatement 형식을 인스턴스에 추가하여 기술한다.

Fig. 9의 9xx로 시작하는 필드는 기관의 독자적인 정보를 기술하기 위해 이용하는 로컬 필드로, MARC 21에서는 사용하지 않아 BIBFRAME 변환 명세가 존재하지 않는다. 그러나 KORMARC에서는 로컬 필드에 대한 대응 필드를 제시하고 있으며 이를 적용하여 900 필드는 700 필드로, 940 필드는 740 필드로, 949 필드는 440 필드로, 950 필드는 020 필드의 입수조건으로 대체 가능하다. 이에 따라 900 필드는 W-contribution-Contribution-agent-Agent-bflc:name00MarcKey 형식으로 변환이 가능하며,

940 필드는 I-title-Title-subtitle-literal 형식으로, 949 필드는 W-title-Title-bflc:title40MarcKey 형태로 기술 가능하다. 950 필드의 가격정보는 I-acquisitionTerms-literal 형태로 변환한다.



Fig. 9. 900/940/949/950 Mapping

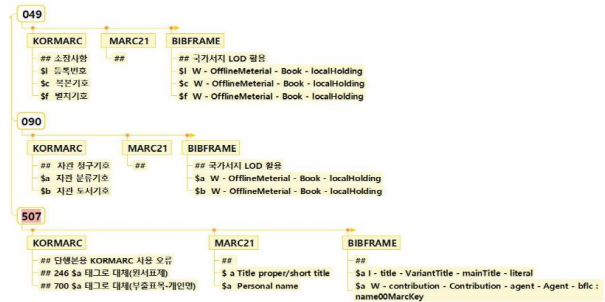


Fig. 10. 049/090/507 Mapping

No attmp convert 분류 Fig. 10의 049 필드는 각 로컬 도서관에서 도서관 부호와 소장 자료의 등록번호, 권.연차 기호, 복본 기호, 별치 기호 등을 기술하는 필드이다. 049 필드는 KORMARC에서만 사용하고 있고 MARC 21에서는 정의되어 있지 않다. 따라서 BIBFRAME 변환 명세를 제공하고 있지 않아 BIBFRAME으로의 변환이 어렵지만, 049 필드는 공공 도서관 서지 데이터에서 출현 빈도가 100%인 매우 중요한 데이터이므로 BIBFRAME 명세에서 공식적으로 기술 방법을 정의해 줄 필요가 있다. 이에 본 연구에서는 국가서지 LOD를 활용하여 변환을 시도한다. 국가서지 LOD를 활용하여 049 필드의 소장 사항 정보도 기술할 수 있도록 코드를 수정하고 BIBFRAME으로 변환을 시도한다. 049 필드의 소장기관 부호, 등록번호, 권.연차기호, 복본기호, 별치기호는 BIBFRAME에서 W-OfflineMeterial-Book-localHolding 형식으로 각각 기술한다.

090 필드는 국립중앙도서관(052 필드)이나 미국 의회 도서관(050 필드)을 제외한 각 도서관 자판 청구기호를 기술하는 필드이다. 090 필드는 MARC 21에서 사용하지 않고

있지만, 로컬에서만 필요한 서지정보이기 때문에 마찬가지로 BIBFRAME 명세가 존재하지 않는다. 090 필드도 출현 빈도가 100%로 매우 높아 BIBFRAME 도입 시 해당 변환 명세가 필요하다. 따라서 090 필드도 국가서지 LOD를 활용하여 자관 청구기호 정보를 기술할 수 있도록 코드를 수정하고 BIBFRAME으로 변환을 시도한다. 090필드의 소장기관 부호, 등록번호, 권.연차기호, 복본기호, 별치기호는 049 필드와 마찬가지로 BIBFRAME에서 W-OfflineMaterial-Book-localHolding 형식으로 각각 기술 가능하다.

507 필드는 그래픽 자료의 축척 주기를 기술하는 필드인데, 노원구립도서관의 분석 데이터에는 필드의 정의와 의미가 다른 정보가 입력되어 있었다. MARC 작업 과정 확인 결과, 납품업체에서 사용하는 MARC 생성 프로그램이 통합 서지용 MARC로 변경되기 전 단행본용 MARC로 세팅되어 있어 단행본용 MARC 정보인 507 원서 주기 필드의 정보를 기록한 것으로 확인되었다. 실제 기록된 단행본용 KORMARC 507 필드의 데이터는 원서 주기(원서 원어 표기, 저자 원어 표기 등)이며 이는 현재 사용하고 있는 통합 서지용 KORMARC의 필드로 변경하여 사용 가능하다. 원서 원어 표기는 246필드로, 저자 원어 표기는 700 필드로 기술 가능하다. 따라서 MARC 21의 246/700 필드를 기준으로 BIBFRAME 변환 명세에 따라 246 필드는 I-title-VariantTitle-mainTitle-literal 형식으로 표현한다. 700 필드는 개인명이 부출 표목으로 채택된 경우에 사용하며 변환 명세에 따라 W-contribution-Contribution-agent-Agent-bflc:name00MarcKey 형식으로 변환 가능하다.

3. Applying the BIBFRAME converter

본 연구에서 제안한 변환 규칙 정의를 검증하기 위해 'marc2bibframe2' 오픈소스 기반의 변환기를 구현하여 실험을 진행한다. 이를 위해 변환에 사용될 KORMARC 서지 데이터 중 적절한 샘플 데이터를 선정한다. 선정된 데이터는 국내외 도서관에서 폭넓게 서지정보를 얻을 수 있는 인기도서 '샬롯의 거미줄(Charlotte's web)'이다.

Table 3. KORMARC to BIBFRAME 2.0 Conversion

BIBFRAME 2.0
<pre>(base) C:\Users\user1\Python\kormarc2bibframe2-master>xlstproc xsl\kormarc2bibframe2.xsl ../Charlotte_test.xml bibframe 변환 명령어 <?xml version="1.0" encoding="UTF-8"?> <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:bf="http://id.loc.gov/ontologies/bibframe/" xmlns:bflc="http://id.loc.gov/ontologies/bflc/" xmlns:madsrdf="http://www.loc.gov/mads/rdf/v1#"> <bf:Work rdf:about="http://example.org/000001317980#Work"> <bf:Document xmlns:nlon="http://lod.nl.go.kr/ontology/"> <nlon:cidp> <rdfs:label>CIP2021012345</rdfs:label> [023 Field \$a] </nlon:cidp> </bf:Document> <bf:Language> <rdf:value rdf:resource="http://id.loc.gov/vocabulary/languages/kor/"> [041 Field \$a] </bf:Language> <bf:Language> </bf:Language> <bf:Language> </bf:Language> <bf:Language> </bf:Language> <bf:part>original</bf:part> </pre>

<pre><rdf:value rdf:resource="http://id.loc.gov/vocabulary/languages/eng/"> [041 Field \$h] </bf:Language> </bf:Language> <nlon:OfflineMaterial xmlns:nlon="http://lod.nl.go.kr/ontology/"> <nlon:Book> <nlon:localHolding>JM99999</nlon:localHolding> [049 Field \$l] <nlon:localHolding>3-35</nlon:localHolding> [049 Field \$v] <nlon:localHolding>3-35</nlon:localHolding> [049 Field \$v] <nlon:localHolding>3</nlon:localHolding> [049 Field \$c] <nlon:localHolding>4</nlon:localHolding> [049 Field \$c] <nlon:localHolding>어린이</nlon:localHolding> [049 Field \$f] <nlon:localHolding>어린이</nlon:localHolding> [049 Field \$f] </nlon:Book> </nlon:OfflineMaterial> <nlon:classification xmlns:nlon="http://lod.nl.go.kr/ontology/"> <nlon:kdc> <nlon:kdcn>808.9</nlon:kdcn> [054 Field \$a] <nlon:kdcn>843</nlon:kdcn> [054 Field \$a] <nlon:editionOfKDC> <rdfs:label>4</rdfs:label> [054 Field \$v2] </nlon:editionOfKDC> </nlon:kdc> </nlon:classification> <nlon:OfflineMaterial xmlns:nlon="http://lod.nl.go.kr/ontology/"> <nlon:Book> <nlon:localHolding>808.9</nlon:localHolding> [090 Field \$a] <nlon:localHolding>시15.1</nlon:localHolding> [090 Field \$b] <nlon:localHolding>3-35</nlon:localHolding> [090 Field \$c] </nlon:Book> </nlon:OfflineMaterial> <bf:Agent> <bf:Agent rdf:about="http://example.org/000001317980#Agent100-13"> <rdf:type rdf:resource="http://id.loc.gov/ontologies/bibframe/Person/"> <bflc:name00MatchKey>와이트, 앨런 브룩스</bflc:name00MatchKey> <bflc:primaryContributorName00MatchKey>와이트, 앨런 브룩스</bflc:primaryContributorName00MatchKey> <rdfs:label>와이트, 앨런 브룩스</rdfs:label> [100 Field \$a] </bf:Agent> <bf:Agent> <bf:Agent rdf:about="http://example.org/000001317980#Agent110-14"> <rdf:type rdf:resource="http://id.loc.gov/ontologies/bibframe/Organization/"> <bflc:name10MatchKey>기본표목, 단체명(TEST)</bflc:name10MatchKey> <bflc:name10MatchKey>110.1 \$a 기본표목, 단체명(TEST)</bflc:name10MatchKey> <bflc:primaryContributorName10MatchKey>기본표목 단체명(TEST)</bflc:primaryContributorName10MatchKey> <rdfs:label>기본표목, 단체명(TEST)</rdfs:label> [110 Field \$a] </bf:Agent> <bf:Agent> <rdfs:label>샬롯의 거미줄</rdfs:label> </bf:Agent> <rdfs:label>샬롯의 거미줄</rdfs:label> <bf:title> <rdfs:label>샬롯의 거미줄</rdfs:label> <bflc:titleSortKey>샬롯의 거미줄</bflc:titleSortKey> <bflc:mainTitle>샬롯의 거미줄</bflc:mainTitle> [245 Field \$a] </bf:title> <bf:hasSeries> </bf:hasSeries> <bf:supplementaryContent> </bf:supplementaryContent> <rdfs:label>서지 등 주기(TEST)</rdfs:label> [504 Field \$a] </bf:supplementaryContent> <bf:supplementaryContent> </bf:supplementaryContent> <bf:supplementaryContent> <rdfs:label>부록 추가(TEST)</rdfs:label> [525 Field \$a] </bf:supplementaryContent> <bf:supplementaryContent> </bf:supplementaryContent> <bf:language> </bf:language> <bf:note> </bf:note> <rdfs:label>영어 원작을 한국어로 번역</rdfs:label> [546 Field \$a] </bf:note> </bf:note> </bf:language> <bf:awards>뉴베리 메달상, 1953</bf:awards> [586 Field \$a] </bf:subject> <bf:Topic rdf:about="http://example.org/000001317980#Topic650-27"> <rdf:type rdf:resource="http://www.loc.gov/mads/rdf/v1/#Topic/"> <rdfs:label>영어 문학(英美文學)</rdfs:label> <madsrdf:authoritativeLabel>영어 문학(英美文學)</madsrdf:authoritativeLabel> [650 Field \$a] </bf:Topic> </bf:subject> </bf:subject> <bf:Topic rdf:about="http://example.org/000001317980#Topic650-28"> <rdf:type rdf:resource="http://www.loc.gov/mads/rdf/v1/#Topic/"> <rdfs:label>창작 동화(創作童話)</rdfs:label> [550 Field \$a] <madsrdf:authoritativeLabel>창작 동화(創作童話)</madsrdf:authoritativeLabel> </bf:Topic> </bf:subject> </bf:subject> <bf:Topic> <rdfs:label>샬롯</rdfs:label> [653 Field \$a] </bf:Topic> </bf:subject> </bf:subject> <bf:Topic> <rdfs:label>거미줄</rdfs:label> [653 Field \$a] </bf:Topic> </bf:subject> </bf:subject> <bf:Topic> <rdfs:label>명미동화</rdfs:label> [653 Field \$a] </bf:Topic> </bf:subject> </bf:contribution> <bf:Contribution> <bf:Agent> <bf:Agent rdf:about="http://example.org/000001317980#Agent700-30"> <rdf:type rdf:resource="http://id.loc.gov/ontologies/bibframe/Person/"> <bflc:name00MatchKey>Williams, Garth, 1912-1996</bflc:name00MatchKey> <bflc:name00MatchKey>7001. \$a Williams, Garth, 1912-1996</bflc:name00MatchKey> <rdfs:label>Williams, Garth, 1912-1996</rdfs:label> [700 Field \$a \$d] </bf:Agent> </bf:Agent> </bf:Contribution> </bf:Contribution> <bf:Agent> <bf:Agent rdf:about="http://example.org/000001317980#Agent700-31"> <rdf:type rdf:resource="http://id.loc.gov/ontologies/bibframe/Person/"> <bflc:name00MatchKey>김화곤</bflc:name00MatchKey> <bflc:name00MatchKey>7001. \$a 김화곤</bflc:name00MatchKey> <rdfs:label>김화곤</rdfs:label> [700 Field \$a] </bf:Agent> </bf:Agent> </bf:Contribution> </bf:Contribution> </bf:Contribution> </bf:Contribution> </nlon:OfflineMaterial> </pre>

2019년 전국 도서관 통계를 참고하여 가장 이용률이 많은 서울시 공공 도서관 중, 장서 보유량이 평균치 이상으로 높은 노원구립도서관의 최신 서지 데이터 5,000개를 이용했다. 분석결과, 방대한 양의 서지 정보 변환 기법 중 25개의 핵심 필드를 도출하여 변환 규칙을 명세하였다.

실제로 사용하는 최신 KORMARC 서지 데이터 분석을 통해 도출한 변환 규칙을 오픈 소스 기반의 변환기 프로그램을 활용하여 KORMARC 전용 변환기를 만들고 국내 최초로 BIBFRAME 변환을 시도했다는 점에 의의가 있으며, KORMARC 데이터 필드의 BIBFRAME 변환 적용 가능성을 확인하였다. KORMARC 서지 데이터의 BIBFRAME 변환을 위해서는 9xx 필드의 로컬 표목 사용을 최소화할 필요가 있으며, KORMARC 데이터 필드의 필드 번호 오기로 인한 변환 오류를 방지해야 한다.

BIBFRAME 2.0의 실제 적용을 위해서는 본 연구에서 추출한 25개의 필드 외에 100여 개가 넘는 데이터 필드에 대한 추가 변환 연구가 필요하며, BIBFRAME 2.0의 공식적인 변환 명세가 없는 로컬 데이터에 대한 연구도 필요하다. 앞으로 BIBFRAME 2.0 자동 변환을 위한 데이터베이스 모델링 및 웹에서 구동하는 변환기에 대한 연구를 후속 연구로 진행하고자 한다.

REFERENCES

- [1] Kyung-ho Lee, and Jung-hyun Kim, "Data Listing Method: Focusing on KORMARC and MARC 21," Daegu: Taeilsa, pp, 186-398, 2016.
- [2] Sung-Sook Lee, and Ji-won Lee, "Study on Conversion between MARC 21 and BIBFRAME 2.0," Journal of Social Science, 29(3), pp, 73-93, Jul 2018.
- [3] <https://librarian.nl.go.kr/LI/contents/L10101000000.do>
- [4] Mihwa Lee, "A Study on the Considerations in Constructing BIBFRAME by Analyzing BIBFRAME 2.0," Journal of Korean Library and Information Science Society, 48(4), pp, 107-127, Dec 2017.
- [5] Suin Yim, and Seungmin Lee, "A Study on the Separation of Descriptive Levels for Enhancing the Applicability of BIBFRAME," JOURNAL OF THE KOREAN SOCIETY FOR LIBRARY AND INFORMATION SCIENCE, 54(3), pp, 165-186, Aug 2020.
- [6] Tae-soo Kim, "Understanding the list," Seoul: Korea Library Association, pp, 165-201, 2008.
- [7] Jung-hyun Kim, "The actuality of the list organization," Daegu: Taeilsa, pp, 73-108, 2018.
- [8] Mihwa Lee, "Analyzing BIBFRAME Cases for the Development of BIBFRAME Application Plans in Korea," Journal of Korean

Library and Information Science Society, 49(2), pp, 59-78, Jun 2018.

- [9] Hye-won Lee, "A Study on Acceptance of BIBFRAME 2.0 for Sharing and Utilization of Bibliographic Data," Korean Library And Information Science Society Summer Conference, 75-90, (), May 2018.
- [10] Xu, Amanda, Hess, Kirk, Akerman, Laura, "From MARC to BIBFRAME 2.0: Crosswalks," Cataloging & classification quarterly, 56(2-3), pp, 224-250, 2018.
- [11] Steele, Thomas D, "What comes next: understanding BIBFRAME," Library hi tech, 37(3), pp, 513-524, 2019
- [12] Zapounidou, Sofia, Sfakakis, Michalis, Papatheodorou, Christos, "Mapping Derivative Relationships from RDA to BIBFRAME 2," Cataloging & classification quarterly, 57(5), pp, 278-308, 2019.
- [13] Samples, Jacquie, Bigelow, Ian, "MARC to BIBFRAME : Converting the PCC to Linked Data," Cataloging & classification quarterly, 58(3-4), pp, 403-417, 2020.
- [14] <https://www.libsta.go.kr>
- [15] <http://www.lod.go..kr>
- [16] Jee-Hyun Rho, "The Current State and Challenges of Linked Data in Library Cataloging," Journal of Korean Library and Information Science Society, 50(3), pp, 71-95, Sep 2019.
- [17] Ok Nam Park, and Jung Sun Oh, "Deployment of BIBFRAME as a New Bibliographic Framework in Linked Data," Journal of the Korean Biblia Society for Library and Information Science, 25(4), pp, 235-263, Nov 2014

Authors



Joo-Yong Kim received the B.S., degree in Library and Information Science from Daejin University, Korea, in 2011, where he is received the in master's degree with the Department of Computer Engineering, in

2021. He is currently a computerized librarian since 2011 in the Public library, Nowon-gu, Seoul, Korea. His research mainly focuses on Linked open data and Database.



Pan-Seop Shin received the B.S., M.S. and Ph.D. degrees in Computer Engineering from Hongik University, Korea, in 1992, 1994 and 2000, respectively. Dr. Shin joined the faculty of the Department of Computer

Engineering at Daejin University, Pocheon-Si, Gyeonggi-Do, Korea, in 2002. He is currently a Professor in the Department of Computer Engineering, Daejin University. He is interested in distributed databases system, data mining and data visualization technology.